

Should a non-rival public good always be provided centrally?*

Nicolas Gravel[†] and Michel Poitevin[‡]

November 6th 2015

Keywords: Federalism, jurisdictions, asymmetric information, equalization, second best, public goods, city mergers

JEL classification codes: D6, H2, H7

Abstract

This paper discusses the problem of optimal design of a jurisdiction structure from the view point of a utilitarian social planner when individuals with identical utility functions for a non-rival public good and private consumption have private information about their contributive capacities. It shows that the superiority of a centralized provision of a non-rival public good over a federal one does not always hold. Specifically, when differences in individuals' contributive capacities are large, it is better to provide the public good in several distinct jurisdictions rather than to pool these jurisdictions into a single one. In the specific situation where individuals have logarithmic utilities, the paper provides a complete characterization of the optimal jurisdiction structure in the two-type case.

"C'est pour unir les avantages divers qui résultent de la grandeur et de la petitesse des nations que le fédératif a été créé." (Alexis de Toqueville)

*We acknowledge with gratitude the financial support of the grant of the French Agency for Scientific Research (ANR contract STRAT-INEQ). We are also indebted, with the usual disclaiming qualification, to Francis Bloch, Massimo Bordignon, Dennis Epple, Pierre Pestieau, Alain Trannoy and Tanguy Van Ypersele for the useful comments that they have made at seminars and conferences where the work leading to this paper has been presented.

[†]Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Centre de la Vieille Charité, 2, rue de la Charité, 13 002 Marseille Cedex, nicolas.gravel@univ-amu.fr.

[‡]Université de Montréal, CIREQ, and CIRANO 3150, rue Jean-Brillant, bureau C-6012, C.P. 6128, succursale Centre-ville Montréal (Québec) H3C 3J7, michel.poitevin@umontreal.ca.

1 Introduction

In many countries, one finds significant regional variation in the bundles of public goods and taxes available to individuals. This is clearly the case in federal countries such as Canada and United States where the provinces and the states have the power to decide the provision of specific public goods (for instance education) and to collect taxes. But this phenomenon is also observed in “unitary” countries such as France or the UK where cities have specific powers in terms of public good production (for instance the financing of primary school infrastructure) and taxation (local taxes). This heterogeneity in the package of public goods and taxes offered to the citizens of a same country is sometimes perceived as the source of unacceptable inequalities. As a result, it is not uncommon to observe attempts made by central authorities to correct these inequalities by means of various cross-jurisdictions equalization payments schemes. But one may wonder why central authorities do not push further this equalizing logic by carrying themselves the task of providing their citizens with the same package of public goods and taxes instead of maintaining these distinct jurisdictions. As the recent North American episodes of city mergers in large agglomerations (Boston, Montreal, Toronto) illustrate, this centralizing solution is sometimes adopted. Yet, the decision to merge many cities into one large agglomeration that is responsible for providing the same public goods and taxes package to all its inhabitants has been received with great skepticism by many. Is this skepticism justified? Is there some argument that can justify cross-citizens heterogeneity in public goods and taxes packages from a normative standpoint or, to put it bluntly, federalism? These are the questions that are addressed in this paper.

These questions are not new. They were underlying the above quote from Alexis de Toqueville and were framed by Wallace Oates (1972) as follows:

“In the absence of cost-savings from the centralized provision of a [local public] good and of interjurisdictional externalities, the level of welfare will always be at least as high (and typically higher) if Pareto-efficient levels of consumption are provided in each jurisdiction than if any single, uniform level of consumption is maintained.”¹

In this paper, we formulate in a precise model Oates’s and Toqueville’s intuition that a federal provision of public goods in separated jurisdictions can be normatively better than a centralized provision “in the absence of cost-saving from the centralized provision”. Specifically, we provide a model in which even when the cost-saving case in favour of a centralized provision of a public good is maximal - namely when the public good is non-rival - it can be optimal to organize its provision in a federal system when the relevant information needed to provide the public good is not available to the social planner.

The formal architecture of the model is as follows. There is a collection of individuals who have the same preferences for one public good and one private good. Each individual has

¹A nice survey of the literature on fiscal federalism is provided by Oates (1999).

an exogenous pecuniary income that is unobserved by the central government. The public good can be provided locally in distinct jurisdictions organized through a federal system or centrally in one grand jurisdiction. In either case, public good provision is financed by taxes in such a way that the budget is balanced within the federation (but not necessarily within a given jurisdiction). The central government chooses the bundles of local public good provision and tax payment - one such a bundle for every jurisdiction - in such a way as to provide incentive for individuals to reveal their willingness to pay for the public good by their locational choice. As in Tiebout (1956), therefore, individuals “vote with their feet” (see also Wildasin (1987)). More specifically, the government chooses bundles of public goods and taxes that maximize a social welfare function under two constraints: 1) a preference revelation constraint (each individual must prefer the combination assigned to him/her by the government to any other) and 2) a budget constraint (the taxes raised in all jurisdictions must be sufficient to finance public goods that are locally provided in all of them).

From a formal point of view, this problem is somewhat reminiscent of the classical Mirrlees optimal income taxation problem (see for instance Mirrlees (1971), Mirrlees (1976), Mirrlees (1986)) with leisure being replaced by a public good. Yet this “replacing” significantly modifies the nature of the problem. In the Mirrlees setting, since leisure is a purely private good - the fact that Bob has 24 hours of leisure per day cannot be used to improve Mary’s utility - there is no intrinsic benefit for the social planner to pool together individuals with different characteristics. Because of this, Mirrleesian optimal income tax schedules typically involve significant separation of workers with different characteristics in order to provide workers with the proper incentives to reveal their type. With public goods, the central planner does benefit from pooling together individuals with different characteristics since a given quantity of a non-rival public good may benefit many individuals at no extra cost. Because of this, the central planner must make a trade-off between the benefit of pooling individuals in order to reduce the cost of public good provision and the welfare cost associated to the provision of the same package of public good and taxation to heterogeneous individuals. The nature of this trade-off determines the optimal level of heterogeneity in public good and taxation packages for a community. It also determines the choice between a centralized provision of the public good within a single “grand” jurisdiction or its provision across separated jurisdictions between which the federal government organizes optimal equalization transfers (see e.g.; Boadway and Flatters (1982), Buchanan (1950), Flatters, Henderson, and Mieszcowski (1974) or Gravel and Poitevin (2006)).

While a few models have examined fiscal federalism issues under asymmetric information in the literature, including those of Aronsson and Blomquist (2008), Bordignon, Manasse, and Tabellini (2001), Breuillé and Gary-Bobo (2007), Cornes and Silva (2000) and Cornes and Silva (2002), none that we are aware of have considered the problem of choosing the appropriate jurisdiction structure – federal or centralized as it may be – for providing a non-rival public good.

The rest of the paper is organized as follows. In the next section, we set up the notation and examine in its full generality the problem of the optimal choice of a jurisdiction struc-

ture under asymmetric information. In section 3, we examine the problem in the (much) more specific setting where there are only two types of individuals, and where these individuals have additively separable and symmetric preferences over the two goods. We provide a complete characterization of the solution of the problem in the important case where the social planner is utilitarian and would like, in the federal solution, to transfer income from the rich jurisdiction to the poor one. As shown in Gravel and Poitevin (2006), this implies indeed that the symmetric and additively separable utility function that represents the individual's preference is in fact logarithmic with respect to both goods. Section 4 provides some conclusions.

2 The general structure of the problem

2.1 Notation

There are $n \geq 2$ individuals taken from a finite set N . Individual i has a monetary income $w_i \in R_+$ and consumes a public good (z) and a private good (x). Individuals are ordered by their income in such a way that $w_i \geq w_{i+1}$ for $i = 1, \dots, n-1$. The public good is non-rival in consumption but is "excludable" in the sense that its consumption may be made contingent upon the fact of belonging to a specific jurisdiction. Specifically, exclusion can be made by partitioning the set N of individuals into pairwise disjoint sets N_j for $j = 1, \dots, l$ for some number $l \in \{1, \dots, n\}$ such that $\bigcup_{j=1}^l N_j = N$. Any set N_j of this partition is interpreted as a jurisdiction and the collection of l such sets $\{N_j\}_{j=1}^l$ is interpreted as a jurisdiction structure. The unique jurisdiction structure that is obtained if $l = 1$ is the grand – or centralized – jurisdiction structure in which all individuals are pooled into one single jurisdiction. All other jurisdiction structures obtained for $l > 1$ are referred to as "federal". An extreme form of a federal structure is obtained when $l = n$ (each individual forms a jurisdiction on his/her own). An allocation of public and private goods for the jurisdiction structure $J = \{N_j\}_{j=1}^l$ is defined as a list $(z_1, \dots, z_l; x_1, \dots, x_n) \in \mathbb{R}_+^{l+n}$ with the interpretation that z_j is the consumption of public good in jurisdiction j (for $j \in \{1, \dots, l\}$) and x_i is the consumption of private good by individual i (for $i \in N$). An allocation of public and private goods $(z_1, \dots, z_l; x_1, \dots, x_n)$ for the jurisdiction structure $J = \{N_j\}_{j=1}^l$ is feasible for that jurisdiction structure if it verifies the federation budget constraint:

$$\sum_{j \in \{1, \dots, l\}} z_j + \sum_{i \in N} x_i \leq \sum_{i \in N} w_i. \quad (1)$$

If $(z_1, \dots, z_l; x_1, \dots, x_n)$ is an allocation of public and private goods that is feasible for the jurisdiction structure $J = \{N_j\}_{j=1}^l$, we denote by $T_i = w_i - x_i$ the tax paid by individual i . An equivalent reformulation of the budget constraint is of course that:

$$\sum_{j \in \{1, \dots, l\}} z_j \leq \sum_{i \in N} T_i$$

(taxes collected must be sufficient to finance the quantities of the public good provided to the citizens).

Individuals convert alternative combinations of private and public goods into utility by the same continuously differentiable, strictly increasing and concave utility function $U : \mathbb{R}_+^2 \rightarrow \mathbb{R}$. The utility function is also assumed to be *super-modular* in the sense of satisfying $U_{xz}(\bar{z}, \bar{x}) \geq 0$ for every bundle $(\bar{z}, \bar{x}) \in \mathbb{R}_+^2$. The second part of the paper also assumes that U is additively separable so that it can be written, for every bundle $(\bar{z}, \bar{x}) \in \mathbb{R}_+^2$, as:

$$U(\bar{z}, \bar{x}) = f(\bar{z}) + h(\bar{x}),$$

for some twice continuously differentiable increasing and concave real-valued functions f and h having both \mathbb{R}_+ as domain. For further use, we denote by V ($V : \mathbb{R}_{++}^3 \rightarrow \mathbb{R}$) the individual's indirect utility function defined as usual by:

$$V(p_z, p_x, R) = \max_{z, x} U(z, x) \text{ subject to } p_z z + p_x x \leq R. \quad (2)$$

We also denote by $z^M(p_z, p_x, R)$ and $x^M(p_z, p_x, R)$ the (Marshallian) demands for public good and private consumption (respectively) when the prices for these two goods are p_z and p_x and when the income of the individual is R . These Marshallian demands are defined, as usual, by the solution to program (2). Given the assumptions imposed on U , it can be seen easily that Marshallian demands and indirect utility are differentiable functions of prices and income that are both decreasing with respect to prices and increasing with respect to income (the two goods are normal if the individuals preferences are represented by a super-modular utility function). We denote by U the class of all direct utility functions that satisfy all these properties and by U_A the subset of U consisting of those functions that are additively separable.

We find convenient to represent individual preferences in the (z, T) space rather than in the (z, x) one. Preferences in (z, T) space, which varies across individuals, are represented by the function $U(\bar{z}, w_i - \bar{T})$. Since the public good is normal, these preferences satisfy the familiar single-crossing property that *the* slope of the indifference curves in (z, T) space is increasing with respect to w_i everywhere, as represented in Figure 1.

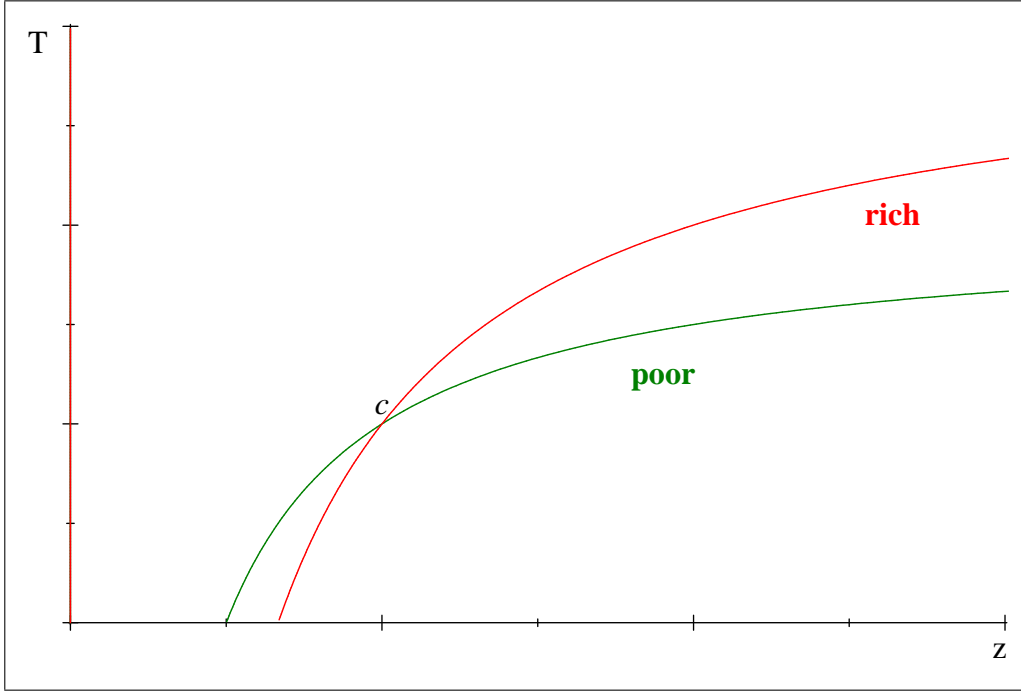


Figure 1: Two indifference curves in (z, T) space.

Alternative allocations of public and private goods are evaluated by a Paretian social planner. For much of the analysis, this planner will actually be *utilitarian*, and will accordingly consider an allocation $(x_1, z_1, \dots, x_n, z_n)$ to be socially better than $(x'_1, z'_1, \dots, x'_n, z'_n)$ if and only if $\sum_i U(z_i, x_i) \geq \sum_i U(z'_i, x'_i)$.

2.2 The choice of a jurisdiction structure under public information

As is well-known from basic public economics, if individuals' utility and income are public information, the problem solved by the social planner is easy. Since the public good is non-rival in consumption, it is a waste to have different individuals consuming different quantities of the public good. For if $(z_1, \dots, z_l; x_1, \dots, x_n)$ is a feasible allocation of private and public goods for a jurisdiction structure $\{N_j\}_{j=1}^l$ with $l > 1$, one can improve everyone's utility by providing everyone with $\bar{z} = \max_{h \in \{1, \dots, l\}} z_h$ units of the public good and $x_i + (\sum_{j \in \{1, \dots, l\}} z_j - \bar{z})/n$ units of the private good. For this reason, the only jurisdiction structure that can be chosen by a Pareto sensitive social planner in a first-best environment is the grand jurisdiction structure associated to $l = 1$. In a first-best world, there is no dispute as to the superiority of a central provision of a non-rival public good over a federal one. There is, of course, a (distributive) need to have different individuals paying different taxes for the public good. More specifically, the utilitarian social planner would choose, in the grand jurisdiction

structure, a distribution of taxes that solves:

$$\max_{(T_1, \dots, T_n)} \sum_{i \in N} U\left(\sum_{i \in N} T_i, w_i - T_i\right). \quad (3)$$

Such a distribution of taxes – call it (T_1^*, \dots, T_n^*) – would satisfy the well-known Samuelson's condition that the sum of the individuals marginal rates of substitution between the private and the public goods equals 1. Since U is concave, and all individuals are consuming the same quantity of the public good, the solution to program (3) would be achieved by a complete equalization of private consumptions. In this ideal first-best world, public and private good consumptions are perfectly equalized, and taxation is individualized by means of Lindahl pricing.

2.3 The choice of a jurisdiction structure under private information

If the information on individuals' characteristics is private, the social planner is no longer able to levy different taxes on individuals who consume the same quantity of the public good because these individuals are indistinguishable from its point of view. The social planner could, however, provide different individuals with different packages of taxes and public good levels based on, say, their place of residence, in order to make them reveal their willingness to pay. Individuals would then "vote with their feet" and choose to live in the jurisdiction offering them their favorite package of public spending and taxes, thus revealing their "type" to the social planner. Of course, separating individuals in different jurisdictions is costly because it requires the use of (much) more resources to provide different levels of the public good in different jurisdictions than what would be needed under a single jurisdiction with a single level of the non-rival public good for all individuals.

We consider herein the case where the private information concerns the individual's income, the only variable that differs across individuals. While the income of a particular individual is unknown to the social planner, we assume that the cumulative density of the income within the population (e.g., the number of individuals who have a income no greater than any real number) is known. With this knowledge, the planner chooses a jurisdiction structure $\{N_j\}_{j=1}^l$ (for some $l \in \{1, \dots, n\}$) and a feasible allocation of the private and the public goods that maximizes its social objective, subject to the constraint that every individual lives in the jurisdiction that provides his/her preferred package of public good and tax. This problem is, in its full generality, complex. Its analysis proceeds in two steps.

In the first step, for a given jurisdiction structure $\{N_j\}_{j=1}^l$ for some $l \in \{1, \dots, n\}$, the social planner solves the program:

$$\max_{z_1, T_1, \dots, z_l, T_l} \sum_{j=1}^l \sum_{i \in N_j} U(z_j, w_i - T_j) \quad (4)$$

subject to the budget constraint:

$$\sum_{j=1}^l z_j \leq \sum_{j=1}^l \#N_j T_j \quad (5)$$

and, for every $j \in \{1, \dots, l\}$, and every $i \in N_j$, the l incentive-compatibility constraints:

$$U(z_j, w_i - T_j) \geq U(z_{j'}, w_i - T_{j'}) \text{ for all } j' \in \{1, \dots, l\}. \quad (6)$$

Let $\Psi(\{N_j\}_{j=1}^l)$ denote the value of the objective function of the social planner at the solution of program (4) under the constraints (5) and (6). The second step of the analysis consists in choosing the jurisdiction structure $\{N_j\}_{j=1}^l$ for some $l \in \{1, \dots, n\}$ that maximizes the value of $\Psi(\{N_j\}_{j=1}^l)$. This second step is a discrete problem since there is only a finite number of different possible partitions of N into jurisdiction structures.

A natural starting point for studying program (4) is to consider the centralized jurisdiction structure. Studying program (4) under the centralized jurisdiction structure is easy because there are no incentive constraints (6) to worry about. In that case program (4) writes (after substituting the budget constraint (5) satisfied at equality into the objective function):

$$\Psi(N) = \max_{T \in [0, w_n]} \sum_{i \in N} U(nT, w_i - T). \quad (7)$$

The necessary (and sufficient by concavity of U) first-order condition for an interior solution T^* of this program can be written as:

$$n \frac{\bar{U}_z^*}{\bar{U}_x^*} = 1, \quad (8)$$

where $\bar{U}_k^* = [\sum_{i \in N} U_k(nT^*, w_i - T^*)]/n$ for $k = z, x$ is the average marginal utility of good k at the optimal choice. Condition (8) looks somewhat like a Samuelson condition. It states that the optimal allocation of public good equalizes n -times the marginal rate of substitution of a *fictitious* individual whose marginal rate of substitution is the ratio of the average marginal utility of the public good over the average marginal utility of the private good. Since the marginal rate of substitution of this fictitious individual has no reason to coincide with that of the average one, the centralized second-best provision of the public good is a priori different from that associated with first-best optimality.

There is, however, an obvious case where condition (8) coincides with the standard Samuelson condition and where, as a result, $\Psi(N)$ is the maximal sum of individual utilities that would obtain in the first-best world where the government had all the relevant information. This case is when the individuals' utility function is quasi-linear (in the private good)

so that it writes, for every z and x , as $U(z, x) = f(z) + x$ for some increasing and concave function f . Indeed with quasi-linear utility, condition (8) writes:

$$n \frac{\partial f(nT^*)}{\partial z} = 1, \tag{9}$$

which is nothing else than the Samuelson's condition associated with quasi-linear utility. Hence, with quasi-linear utility, it is possible to achieve a first-best allocation of private and public goods by pooling everybody in the same jurisdiction. Because of this, in the quasi-linear case, one has $\Psi(N) \geq \Psi(\{N_j\}_{j=1}^l)$ for every jurisdiction structure $\{N_j\}_{j=1}^l$ that the central planner could consider. In a quasi-linear world, it would never be optimal to create more than one jurisdiction for providing a non-rival public good.

If the assumption of quasi-linearity is relaxed, it is not clear that the centralized structure dominates the federal one. The question is thus: under which conditions is a centralized jurisdiction structure optimal? This question is difficult to answer in general. For one thing, the two-step program described above is hard to solve because if the number of individuals with different income levels is large, then so is the number of jurisdiction structures that are to be compared. The single-crossing property simplifies however program (4). Among other things, it guarantees that all jurisdictions of a given federation that satisfy incentive constraints (6) will be consecutive (e.g., see Greenberg (1983)) and, therefore, be such that for any two jurisdictions with distinct per capita income, the richest individual of the poorer jurisdiction is weakly poorer than the poorest individual of the richer one. Yet, significant as it is, this simplification leaves the analysis of program (4) in its full generality still quite difficult. One of the difficulty lies in handling the incentive constraints that may or may not be binding. Because of this complexity, we focus attention herein on the simpler problem of characterizing the optimal jurisdiction structure when there are only two types of individuals.

3 The choice of a jurisdiction structure with two types of individuals

We assume accordingly that the n individuals can be split into two types: n_1 individuals of type 1 (the "rich" with income w_1) and n_2 individuals of type 2 (the "poor" with income w_2). In this setting the choice of the optimal jurisdiction structure made by the social planner amounts to comparing the value of its objective function in the centralized jurisdiction structure with that in a federation where individuals are split into two jurisdictions: one inhabited by the poor and the other, by the rich. It is clear indeed that, beside this segregated federation in which all rich individuals live in a jurisdiction and all poor in another, there is no partition of the set of individuals into (more than one) disjoint subsets that can maximize a Pareto-inclusive objective while satisfying the incentive-compatibility constraints. We denote by (N_1, N_2) the federal jurisdiction structure in which the rich and the poor live in two distinct

jurisdictions and by N the centralized jurisdiction structure in which all individuals live in the same jurisdiction.

3.1 The two-type case with general utility functions

Consider first the centralized jurisdiction structure. Let us examine the utility possibility set of this structure. The Pareto frontier of this set is easy to characterize. It is a curve lying between two points: One where each poor gets his "ideal utility" associated to his favorite tax and the rich gets the utility associated with the poor's favorite tax and the other extreme situation where the rich gets her ideal utility and the poor gets the utility level associated to the fact of paying the most preferred tax of the rich.

The favorite tax rate T_i^* of a type i individual is the solution of the program:

$$\max_{T \in [0, w_2]} U(nT, w_i - T).$$

for $i = 1, 2$ It is therefore defined by $T_i^* = \min[z^M(1/n, 1, w_2)/n, w_2]$. Hence the poor's ideal utility level is $V(1/n, 1, w_2)$ while the rich's ideal utility level, denoted U_{SB1}^{1*} with the subscript standing for the "second-best 1-jurisdiction") is defined by:

$$U_{SB1}^{1*} = \begin{cases} V(1/n, 1, w_1) & \text{if } z^M(1/n, 1, w_1)/n \leq w_2 \\ U(nw_2, w_1 - w_2) & \text{otherwise.} \end{cases}$$

It is immediate to see that, if the public good is normal, $T_1^* > T_2^*$ and that:

$$U_{SB1}^{1*} > U(z^M(1/n, 1, w_2), w_1 - T_2^*) > V(1/n, 1, w_2) > U(nT_1^*, w_2 - T_1^*).$$

Hence, the two extreme points $(U(z^M(1/n, 1, w_2), w_1 - T_2^*), V(1/n, 1, w_2))$ and $(U_{SB1}^{1*}, U(nT_1^*, w_2 - T_1^*))$ of the Pareto frontier lie in the area where the utility of the rich is larger than that of the poor. It is therefore impossible to be egalitarian in this second-best world with one jurisdiction. Between its extreme points, the Pareto frontier is defined by the function $\Theta : \mathbb{R} \rightarrow \mathbb{R}$ with $\Theta(u_1) = U(nT^{SB}(u_1), w_2 - T^{SB}(u_1))$ where the function T^{SB} solves implicitly the equality:

$$U(nT, w_1 - T) = u_1.$$

for all $T \in [T_2^*, T_1^*]$. It is easy to verify by usual implicit function arguments that the function Θ is decreasing and concave on its domain.

This frontier is illustrated in Figure 2 for a situation where $U_{SB1}^{1*} = V(1/n, 1, w_1)$ (the ideal tax of the rich is less than the poor's income).

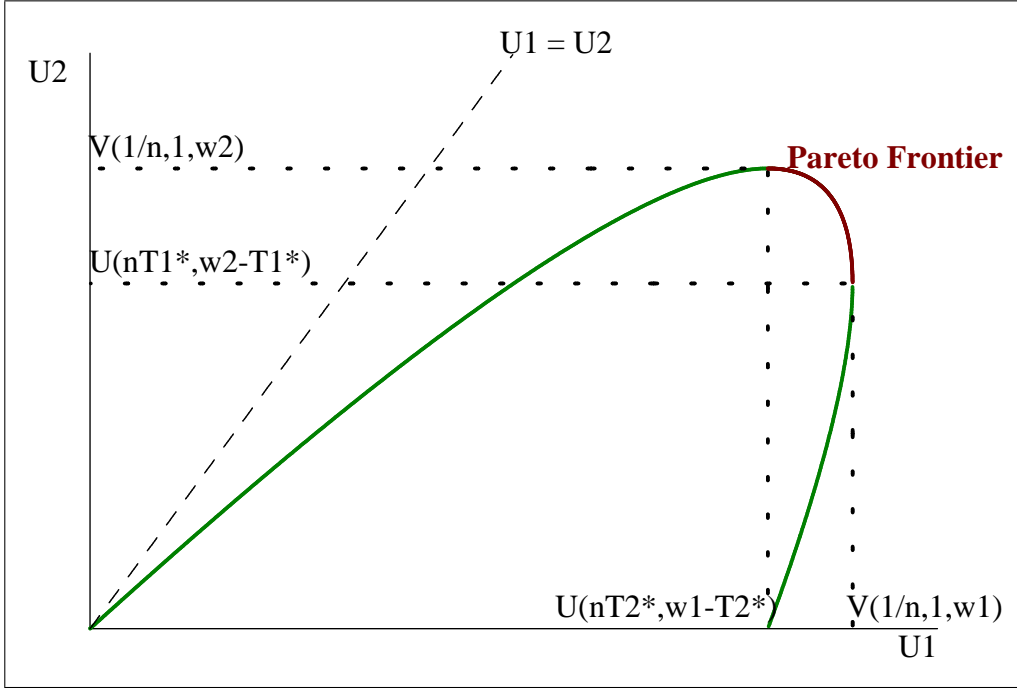


Figure 2: Utility possibility set of a second best centralized jurisdiction structure.

Notice that T_2^* is the solution that would be selected by an extremely egalitarian leximin or a maximin social planner.

We now turn to the federal jurisdiction structure (N_1, N_2) in which the central government separates the two types into two different jurisdictions by solving program (4) for that case. A first possibility that can arise is that none of the two IC constraints (6) of this program binds. If this is the case, one can solve program (4) by ignoring these constraints. This program, studied for an arbitrary number of jurisdictions in Gravel and Poitevin (2006), describes how a utilitarian central planner would design optimal equalization payments in a federal system with an immobile population. Let $B(n_1, n_2, w_1, w_2) = \{(z_1, T_1, z_2, T_2) \in R^4 : n_1 T_1 + n_2 T_2 \geq z_1 + z_2 \text{ and } T_i \leq w_i \text{ for } i = 1, 2\}$ denote the set of two-jurisdiction packages of public goods and taxes that are feasible for a two-jurisdiction segregated federation notwithstanding the IC constraints. It is easy to establish the following lemma, proved, as all formal results of this paper, in the Appendix.

Lemma 1 *Let U be a utility function in U . Then $(z_1^*, T_1^*, z_2^*, T_2^*)$ is Pareto-efficient in the set $B(n_1, n_2, w_1, w_2)$ if and only if there exists $s_i^* \in [-w_i, +\infty)$ satisfying $z_i^* = z^M(1/n_i, 1, w_i + s_i^*)$ and $T_i^* = w_i - x^M(1/n_i, 1, w_i + s_i^*)$ for $i = 1, 2$, and $n_1 s_1^* + n_2 s_2^* = 0$.*

This lemma says that any Pareto-efficient allocation of public goods and tax burdens in a two-jurisdiction federation – ignoring the incentive compatibility constraints – can be

thought of as resulting from a two-step procedure: a first step in which a federal government selects a pair of per capita net equalization subsidies s_1 and s_2 (one such subsidy for every jurisdiction, with aggregate subsidies summing to 0) and a second step in which each individual allocates his/her income – increased by the subsidy received – between public and private good expenditures assuming that he/she faces a price of the public good given by the inverse of the population size of his/her jurisdiction of residence.

An efficient federal provision of public goods and tax burdens in this sense (that satisfies the incentive constraints (6)) is depicted on Figure 3.

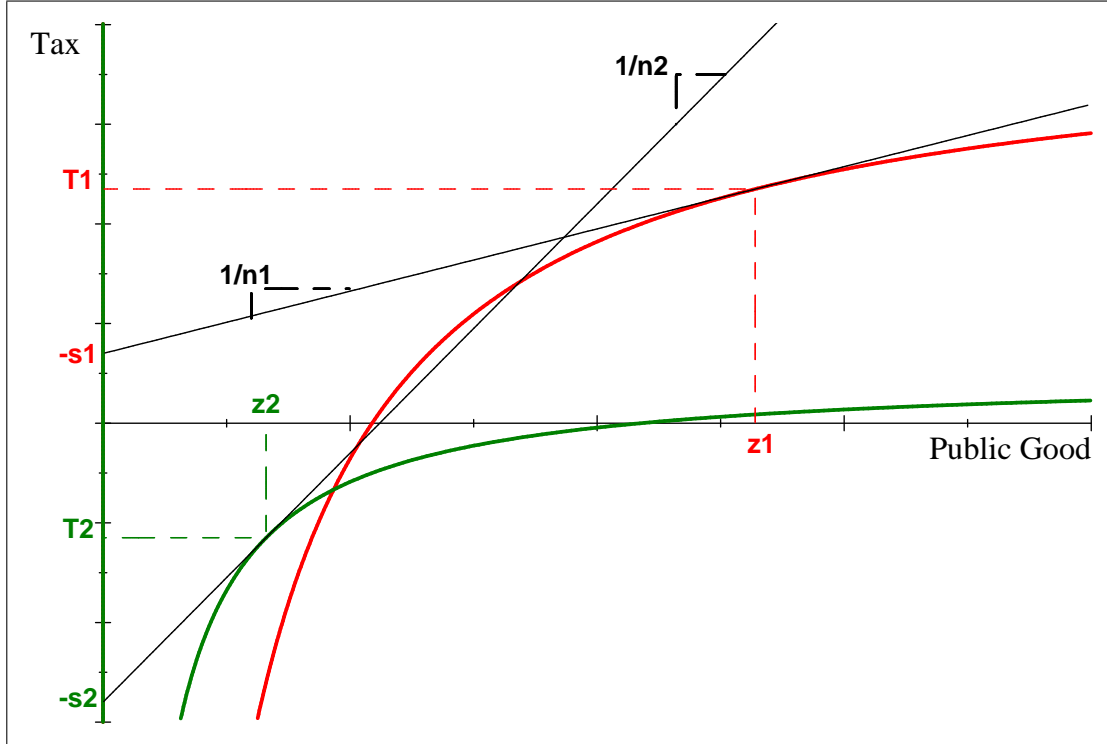


Figure 3: An efficient federal allocation that satisfies the incentive constraints.

Notice that the two subsidies depicted in Figure 3 are progressive. In a two-type setting, the net subsidy "received" by the rich individual is negative and the net subsidy received by the poor one is positive. The progressivity of subsidies would seem to be a somewhat natural feature of a federal organization. For it indicates that "rich" jurisdictions transfer part of their tax revenues to poorer ones in order to provide comparable access to public goods of the type often required in federal countries like Canada.² It happens however that

²"Parliament and the Government of Canada are committed to the principle of making equalization payments to ensure that provincial governments have sufficient revenues to provide reasonably comparable levels of public services at reasonably comparable levels of taxation." (Subsection 36(2) of the Constitution Act, 1982)

the progressivity of subsidy is not a genuine characteristic of utilitarian optimal equalization payments in a federation. As shown in Gravel and Poitevin (2006), it is indeed possible that the subsidies chosen by a utilitarian social planner be regressive in the sense of being increasing with respect to individual income. Regressivity may arise whenever (and only whenever) the objective function of the social planner is not additively separable between individual income and the price of public good. If the social planner is utilitarian, then a necessary and sufficient condition for avoiding this possible regressivity is for the indirect utility function to be itself additively separable between the price of the public good and the income of the individual. In Gravel and Poitevin (2006), it was further shown that if the individual indirect utility function results from an additively separable individual direct utility functions, then it can be additively separable between the price of the public good and the income if and only if the direct utility function is either quasi-linear (in the private good) or logarithmic with respect to the public good. As mentioned above, quasi-linear utility makes the problem studied in this paper trivial. Hence, if we want to limit our attention to additively separable utility functions while being sure that the subsidies chosen by the utilitarian social planner (in the absence of incentive-compatibility constraints) are progressive, we must assume that the additively separable function is in fact logarithmic with respect to the public good.

Beside its empirical plausibility, considering (optimally chosen) progressive per capita subsidies in federations with immobile individuals also simplifies the handling of incentive compatibility constraints in the study of the optimal jurisdiction structure. For it can be seen that if a pair of public good and tax packages that is Pareto efficient in the federation for a progressive subsidies scheme violates the incentive constraint of the poor, then the centralized jurisdiction structure Pareto dominates the federal one. We state this fact formally as follows.

Proposition 1 *Let U be a utility function in U and assume that $(z_1^*, T_1^*, z_2^*, T_2^*)$ is efficient in $B(n_1, n_2, w_1, w_2)$ with respect to a scheme of subsidies s_1 and s_2 as per Lemma 1 satisfying $s_1 < 0 < s_2$. Then, if $U(z_2^*, w_2 - T_2^*) \leq U(z_1^*, w_2 - T_1^*)$ (incentive compatibility constraint of the poor is binding or violated), $\Psi(N) > \Psi(N_1, N_2)$.*

Simple as it is, this lemma facilitates the analysis of the two-type case by restricting attention, in the study of the optimal federal provision of the public good, to the incentive compatibility constraint of type-1 individuals. Of course, the progressivity of the equalization subsidies that supports an efficient federal provision of public goods as per Lemma 1 is crucial for Proposition 1. Yet, if we take this progressivity as granted, the only relevant possibility other than that when no IC is binding when analyzing program (4) in the two-type case is to solve:

$$\begin{aligned} \max_{z_1, T_1, z_2, T_2} \quad & n_1 U(z_1, w_1 - T_1) + n_2 U(z_2, w_2 - T_2) \\ \text{s.t.} \quad & z_1 + z_2 \leq n_1 T_1 + n_2 T_2 \\ & U(z_1, w_1 - T_1) \geq U(z_2, w_1 - T_2). \end{aligned} \tag{10}$$

Two general remarks can be made about this program. A *first* one is that its solution entails the well-known “no distortion at the top” property that the bundle of public and private goods consumed by a rich individual in a federal system where her incentive constraint is binding is the bundle that she would like to consume in such a federal system without incentive constraint if she was facing a price of public good of $1/n_1$ and had a income of $w_1 - s$ for some per capita subsidy s transferred to the poor jurisdiction. A formal way to state this is that $(\hat{z}_1, \hat{T}_1, \hat{z}_2, \hat{T}_2)$ is the solution of program (10) if and only if $\hat{z}_1 = z^M(1/n_1, 1, w_1 - \hat{s})$ and $\hat{T}_1 = w_1 - x^M(1/n_1, 1, w_1 - \hat{s})$ for the per capita subsidy \hat{s} that solves, along with (\hat{z}_2, \hat{T}_2) , the program:

$$\begin{aligned} \max_{s, T_2, z_2} \quad & n_1 V(1/n_1, 1, w_1 - s) + n_2 U(z_2, w_2 - T_2) \\ \text{s.t.} \quad & n_1 s + n_2 T_2 \geq z_2 \\ & V(1/n_1, 1, w_1 - s) \geq U(z_2, w_1 - T_2). \end{aligned} \tag{11}$$

The *second* remark concerns the relation between the per capita subsidy given by rich individuals to poor ones when the incentive constraint of the rich individual binds (Program (11)) and the subsidy when this incentive constraint does not bind (Lemma 1). As shown in the following proposition, the per capita subsidy given by the rich jurisdiction to the poor one will always be smaller when the incentive constraint binds than when the incentive constraint does not bind. Hence, as in the Mirrleesian optimal taxation literature, giving incentives to the rich to stay in their jurisdiction somewhat mitigates the equalizing propensity of the social planner as compared to what it would do if the incentives of the rich were not constraining.

Proposition 2 *Let U be a utility function in U and s^* be the solution of the program:*

$$\max_s n_1 V(1/n_1, 1, w_1 - s) + n_2 V(1/n_2, 1, w_2 + n_1 s/n_2). \tag{12}$$

Let $(\hat{s}, \hat{z}_2, \hat{T}_2)$ be the solution of program (11). Then $\hat{s} \leq s^$ and $\hat{z}_2 \leq z_2^M(1/n_2, 1, w_2 + n_1 s^*/n_2)$.*

The reason for the fact that the subsidy from the rich to the poor in a federation is smaller when the incentive constraint binds than when it does not is clear. When the incentive constraint does not bind, the subsidy serves to some extent the purpose of equalizing the marginal utility of income of both types of individuals. It acts therefore as a redistributing device. When the incentive compatibility constraint is binding, the central planner must cool down its redistributive propensity and must leave rich individuals with a higher income and a lower marginal utility of income than that of the poor. This is achieved with a lower subsidy. Note that the cost of redistribution is in terms of distortion of the amount of public good consumed by the poor individual. If the two goods are normal, this means that the rich individual will consume more of the public good when the constraint is binding than

when it is not. As for the poor individual, he will consume less of the public good when the constraint is binding. In general, it is not possible to determine whether the tax paid by the poor individual is larger or smaller when the incentive constraint binds as compared to the situation where it does not bind. On the one hand, less resources are needed since less public good is provided. On the other hand, the subsidy received from the rich is smaller so that more taxes may need to be raised.

We have so far established two noticeable general results. First, we have shown that the central structure always dominates the federal one when the incentive constraint of the poor is binding (Proposition 1). We have also demonstrated that the per capita subsidy given to the poor is lower when the incentive constraint of the rich is binding than when it is not (Proposition 2). In order to get more specific results, we consider in the next section a specialization of our model to additively separable utility functions.

3.2 The two-type case with logarithmic utility functions

In what follows, we compare the federal and the centralized provisions of the public good with private information in the two-type case from the view point of a utilitarian social planner when the individual's utility function writes:

$$U(z, x) = \ln z + \ln x. \quad (13)$$

The relevance of this specification has already been discussed in the preceding section. Another advantage of this specification is that, for the most part, it makes the choice between a federal and centralized structure depending upon only two parameters: the ratio of the high income over the small one (the interquartile ratio) and the ratio of the number of rich over the number of poor (demographic ratio). As will be seen, this two-dimensional representation is quite useful for identifying the set of parameters that determine the social planner's preference for the federal structure over the centralized one. Among other things, it enables a nice two-dimensional geometric depiction of the situation.

We start by finding the provision of public good and tax that the planner would choose in the centralized jurisdiction structure as per program (7). The first-order condition of program (7) writes:

$$n_1 U_z(z^{sb1}, w_1 - T^{sb1}) + n_2 U_z(z^{sb1}, w_2 - T^{sb1}) = \frac{n_1 U_c(z^{sb1}, w_1 - T^{sb1}) + n_2 U_c(z^{sb1}, w_2 - T^{sb1})}{n}.$$

Applying this to the logarithmic utility function defined by (13) yields (after lengthy manip-

ulations done with the precious help of *Mathematica*):³

$$\begin{aligned} z^{sb1} &= \frac{n_1 w_1 + n_2 w_2 + 2n_1 w_2 + 2n_2 w_1 - g(n_1, n_2, w_1, w_2)}{4}, \\ T^{sb1} &= \frac{z^{sb1}}{(n_1 + n_2)}, \end{aligned}$$

where:

$$g(n_1, n_2, w_1, w_2) = (-8(n_1 + n_2)^2 w_1 w_2 + (n_2(2w_1 + w_2) + n_1(w_1 + 2w_2))^2)^{1/2},$$

and where the superscript *sb1* refers to the 1-jurisdiction second-best allocation.

We now consider the optimal allocation of public good and taxes that the utilitarian social planner would choose in the federal system in which the two types of individuals are separated into two distinct jurisdictions. Again, this amounts to analyzing program (4) for the specific logarithmic utility function of equation (13).

As discussed above, we start this analysis by first considering the case where the two incentive constraints are satisfied and where the social planner solves:

$$\max_{z_1, T_1, z_2, T_2} n_1 [\ln z_1 + \ln(w_1 - T_1)] + n_2 [\ln z_2 + \ln(w_2 - T_2)] \quad \text{s.t.} \quad z_1 + z_2 \leq n_1 T_1 + n_2 T_2. \quad (14)$$

It is not difficult to show that the solution of program (14) is:

$$\begin{aligned} z_1^{fb2} &= \frac{n_1(n_1 w_1 + n_2 w_2)}{2(n_1 + n_2)}, \\ z_2^{fb2} &= \frac{n_2(n_1 w_1 + n_2 w_2)}{2(n_1 + n_2)}, \\ T_1^{fb2} &= \frac{n_1 w_1 - n_2 w_2 + 2n_2 w_1}{2(n_1 + n_2)}, \\ T_2^{fb2} &= \frac{n_2 w_2 - n_1 w_1 + 2n_1 w_2}{2(n_1 + n_2)}, \end{aligned}$$

where the superscript *fb2* refers to the 2-jurisdiction first-best allocation in which incentive constraints are ignored. *Two* things should be noticed about this solution. *First*, it implies that private good consumption is equalized across jurisdictions because the individual's income (net of the subsidy received) is equalized across jurisdictions. Notice that the equalization of income results from the fact that the indirect utility function is additively separable in the income and the price of the public good. However, the quantity of the public good consumed in the two jurisdictions differs because the price of the public good (equal to the inverse of the population size) differs. The second thing that can be observed is that

³All computations performed on *Mathematica* are available upon request.

the level of public good in one jurisdiction is increasing with respect to that jurisdiction's population and decreasing with respect to the other jurisdiction's population.

We can compute the per capita subsidy (as per Lemma 1) of type i that goes from the rich to the poor because of the progressivity associated with the considered preferences. This subsidy, denoted s_i^{fb2} , is the per capita difference between tax revenues and public good expenditure in jurisdiction i

$$s_i^{fb2} \equiv \frac{T_1^{fb2} - z_i^{fb2}}{n_i} = \frac{n_j(w_1 - w_2)}{n_1 + n_2}. \quad (15)$$

for $i, j = 1, 2$ and $j \neq i$. The subsidy is monotonically increasing with respect to the income difference between the two types of individuals. The subsidy is also increasing with respect to the population size of each jurisdiction. When population increases in one jurisdiction, the demand for public good increases. For the assumed utility function, the increase in public good is exactly financed by the increase in population so that per capita taxes do not have to increase. When n_2 increases, the per capita subsidy to the poor decreases so that the marginal utility of income for the poor increases. To restore the equality of marginal utility of income across jurisdictions, the subsidy has to increase. When n_1 increases, the per capita subsidy from the rich decreases so that their marginal utility of income has increased. The subsidy then has to increase for optimal redistribution.

We now consider the incentive compatibility constraints. Since we do not need to worry about the incentive constraint of the poor type in so far as the comparison of the federal and the centralized provision of the public good is concerned thanks to Proposition 1, we restrict our attention to the program:

$$\begin{aligned} \max_{z_1, T_1, z_2, T_2} \quad & n_1[\ln(w_1 - T_1) + \ln z_1] + n_2[\ln(w_2 - T_2) + \ln(z_2)] \\ \text{s.t.} \quad & z_1 + z_2 \leq n_1 T_1 + n_2 T_2 \\ & \ln(w_1 - T_1) + \ln z_1 \geq \ln(w_1 - T_2) + \ln z_2. \end{aligned} \quad (16)$$

If the incentive constraint of the rich does not bind at the solution of program (16), then the solution of this program is precisely that of program (14) above. Hence, we start by evaluating the incentive constraint of the rich at the solution of program (14). We can show (again with the help of *Mathematica*) that:

$$\ln(w_1 - T_1^{fb2}) + \ln z_1^{fb2} \geq \ln(w_1 - T_2^{fb2}) + \ln z_2^{fb2}$$

is equivalent to:

$$\frac{n_1^2 w_1 + n_2^2(w_2 - 2w_1) + 3n_1 n_2(w_2 - w_1)}{n_2(3n_1 w_1 + 2n_2 w_1 - 2n_1 w_2 - n_2 w_2)} \geq 0. \quad (17)$$

Since the denominator is positive, only the numerator matters for the sign of the expression. We can restrict the number of parameters by replacing n_1 by $a * n_2$ and w_1 by $b * w_2$ where:

$$a = n_1/n_2$$

is the demographic ratio (of the number of rich over the number of poor) and

$$b = w_1/w_2$$

is the interquartile ratio (the ratio of the highest income over the smallest income in this two-type world). By assumption $a > 0$ and $b > 1$. Condition (17) can then be equivalently rewritten as:

$$IC_1(a, b) \equiv a^2b - 3(b-1)a + 1 - 2b \geq 0. \quad (18)$$

When inequality (18) holds, the incentive constraint for the rich individual is satisfied at the $fb2$ allocation. While, thanks to Proposition 1, this is not needed for the comparison of the federal and the centralized structures, we can also write down the inequality that corresponds to the incentive constraint of the poor. This inequality can be shown to be equivalent to:

$$IC_2(a, b) \equiv a^2(b-2) + 3(b-1)a + 1 \geq 0. \quad (19)$$

Figure 4 provides a graphical representation of these two incentive constraints in (a, b) -space.

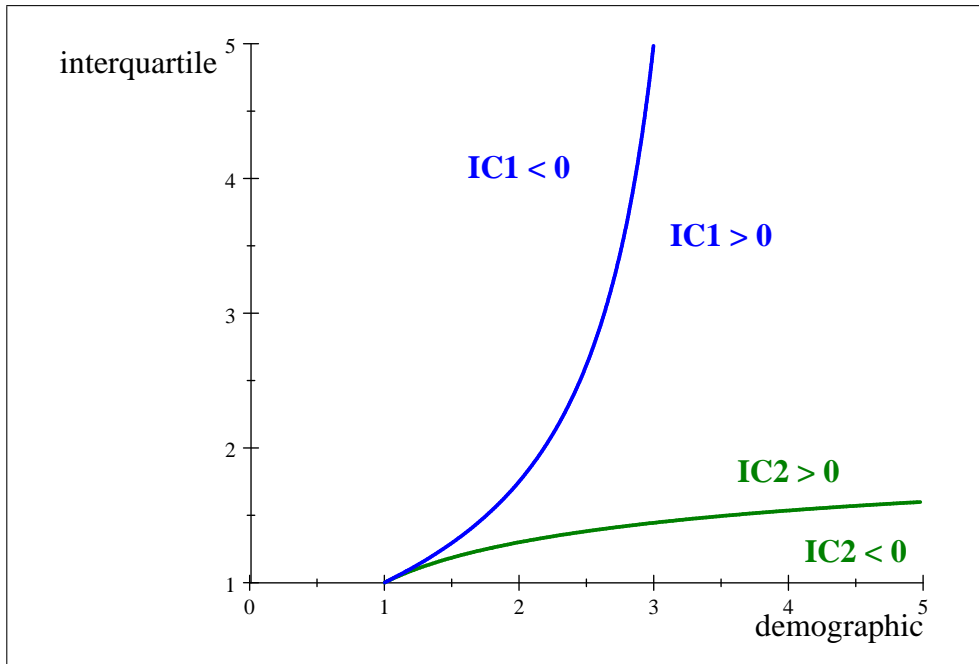


Figure 4: Incentive constraints in (a, b) -space.

As can be seen, the curves IC_1 and IC_2 intersect at $a = b = 1$, which is to be expected. When $b = 1$, all individuals have the same income. Hence, if they live in two separated jurisdictions, they will only differ by their consumption of the public good, which depends on the population of their jurisdiction. When $a = 1$, both jurisdictions have the same population,

hence the same consumption of the public good. Both incentive-constraints are therefore satisfied at this point.

For all values of $a < 1$, the incentive-constraint of the rich individual is violated at the solution of program (14). That is, rich individuals always have an incentive to move from their jurisdiction to that of the poor when there are more poor than rich. Indeed, when the number of rich is lower than that of the poor, the rich face a larger price of public good and, as a result, consume less public good than the poor. Since the poor also pay less tax than the rich, the rich have incentive to join the poor jurisdiction. In the case where the number of rich is smaller than that of the poor, the social planner will face a binding incentive constraint of the rich if it chooses a federal structure. Will it? In order to answer this question, we first compare the centralized and the federal structures when the incentive constraint of the rich is satisfied at allocation $fb2$.

The social welfare at the second-best allocation with one jurisdiction is:

$$SW^{sb1} = n_1 U(z^{sb1}, w_1 - T^{sb1}) + n_2 U(z^{sb1}, w_2 - T^{sb1}). \quad (20)$$

while the social welfare at the two-jurisdiction federal allocation without incentive constraints is:

$$SW^{fb2} = n_1 U(z_1^{fb2}, w_1 - T_1^{fb2}) + n_2 U(z_2^{fb2}, w_2 - T_2^{fb2}). \quad (21)$$

After some manipulations (again performed by *Mathematica*), we can show that:

$$SW^{fb2} - SW^{sb1} \propto 4^{1+a} a^a (1 + ab)^{2+2a} - (1 + a)^{1+a} (1 + 2b + a(2 + b) - h(a, b))^{1+a} \times \\ (3 + 2a - 2b - ab + h(a, b)) (-1 - 2a + 2b + 3ab + h(a, b))^a$$

where $a = n_1/n_2$, $b = w_1/w_2$, and:

$$h(a, b) = \sqrt{-8(1 + a)^2 b + (1 + 2b + a(2 + b))^2}.$$

The sign of $SW^{fb2} - SW^{sb1}$ is the same as the sign of the expression on the right-hand side. This is quite a messy expression, but it does depend only upon the demographic and the interquartile ratios. This expression is plotted (in red color) in (a, b) -space in Figure 5.

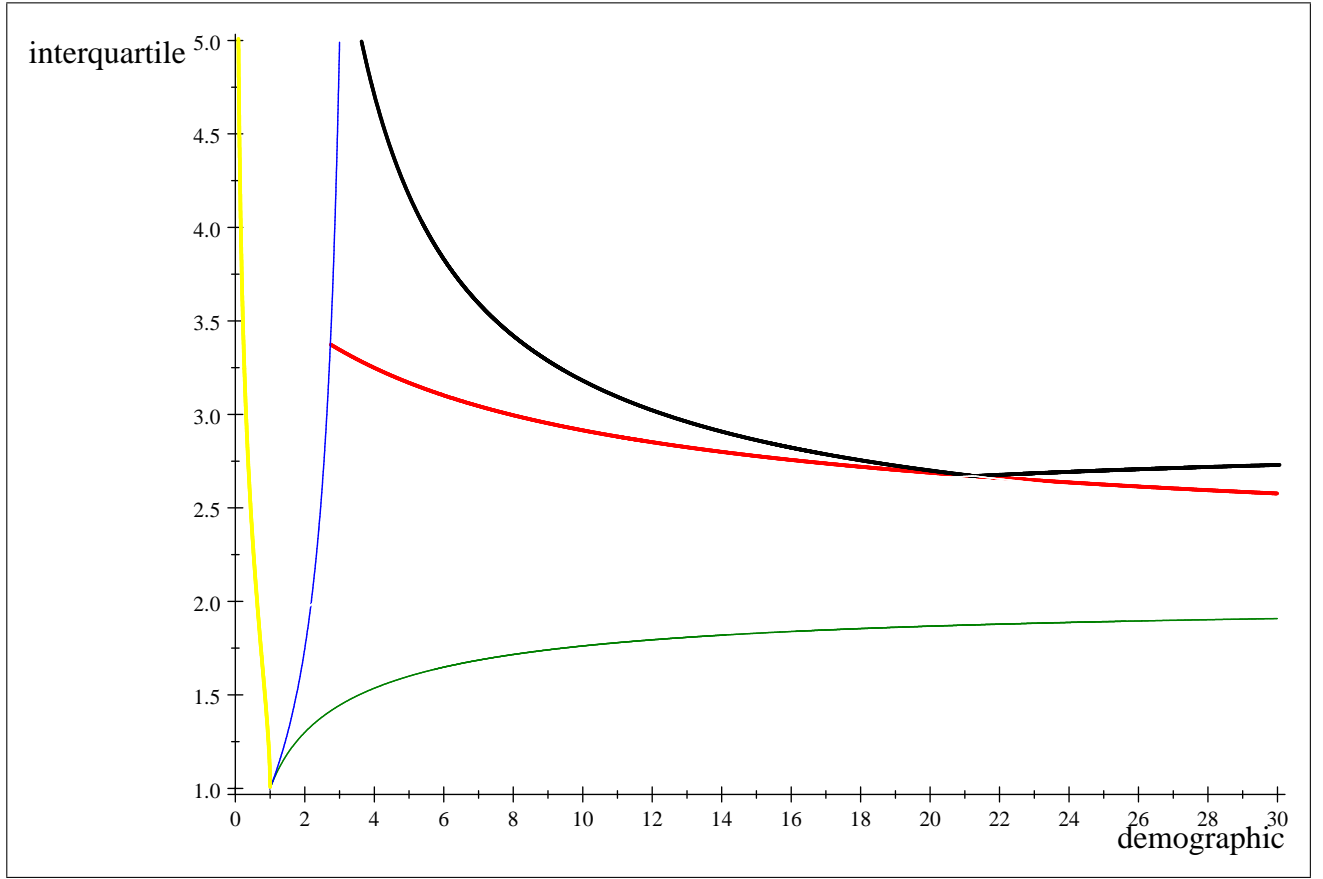


Figure 5: Distributional parameters that determine the choice of the optimal jurisdiction structure.

Over the region where both incentive constraints (in blue and in green as in Figure 4) are satisfied – that is, when the allocation $fb2$ is the solution to the second-best problem – the set of combinations of values for the parameters (a, b) that lie above the red curve are those for which the utilitarian social planner prefers the federal structure to the centralized one. As can be seen, this preference arises when income disparities, as measured by the interquartile ratio w_1/w_2 , are relatively large (so that there is a strong motive for income redistribution and adjustment of the private and public good provision to the heterogeneity of individuals willingness to pay for the public good). Analogously, the larger is the relative number of rich, the more likely it is that the federal solution be favoured by the utilitarian planner. When there are more rich, it is less costly to redistribute to the poor since less income has to be taken away from each rich individual in order to achieve a given amount of redistribution. In addition, Figure 5 illustrates the result of Proposition 1 since the red curve on the North East of which the federal solution dominates the centralized one never intercepts the boundary of the zone where the incentive constraint of the poor (in green) is violated.

Figure 5 also uses the Pareto criterion to compare the two structures. The area above the black curve indicates indeed the set of parameters values for which the utilitarian optimal federal structure is unanimously preferred to the utilitarian optimal centralized one. The area above the black curve is simply the set of all parameter values for which:

$$U(z^{sb1}, w_1 - T^{sb1}) \leq U(z_1^{fb2}, w_1 - T_1^{fb2})$$

$$U(z^{sb1}, w_2 - T^{sb1}) \leq U(z_2^{fb2}, w_2 - T_2^{fb2}),$$

so that both types of individuals prefer the *fb2* allocation to the *sb1*.

Unfortunately, the two-dimensional feature of the analysis in (a, b) -space is lost when the incentive constraint of the rich is binding at the solution of program (16). For we cannot extend the red curve in Figure 5 to the northwest of the IC1 (blue) curve. In this case, the choice between the two structures depends upon the four parameters (n_1, n_2, w_1, w_2) , and not only upon the ratios $a = n_1/n_2$ and $b = w_1/w_2$. We can nonetheless state results about the set of values of the parameters that would lead the utilitarian social planner to favor the federal system over the centralized one. The key to this part of the analysis is the study of the subsidy given by the rich to the poor that solves Program (11). As shown in Proposition 2, this subsidy is smaller when the IC constraint of the rich is binding at the optimal allocation in a federal structure than when it is not binding. It is even possible for this subsidy to be negative when the IC constraint of the rich is binding.

Indeed, consider fixing s in program (16) and solving the program for z_2 and T_2 using the two constraints for that s . Denote this partial solution $z_2(s)$ and $T_2(s)$. Now, the optimal subsidy \hat{s} solves:

$$\max_s n_1 V(1/n_1, 1, w_1 - s) + n_2 U(z_2(s), w_2 - T_2(s)). \quad (22)$$

and satisfies therefore the first-order condition:

$$-n_1 V_R(1/n_1, 1, w_1 - \hat{s}) + n_2 \left(\frac{1}{z_2(\hat{s})} \frac{dz_2(\hat{s})}{ds} - \frac{1}{w_2 - T_2(\hat{s})} \frac{dT_2(\hat{s})}{ds} \right) = 0.$$

There are values of (w_1, w_2, n_1, n_2) for which this condition is solved for $\hat{s} = 0$. It is not difficult (thanks to *Mathematica*) to see that the set of parameters for which this happens depends only upon a and b . This set is in fact described by the equality:

$$b = \frac{-\sqrt{1-a} - 2a}{-2a + a\sqrt{1-a}}. \quad (23)$$

On Figure 5, we have plotted (in yellow) the curve described by equation (23). It is easy to see that b goes to infinity when a becomes negligible. For all values of a and b to the left (right) of this line, the optimal subsidy is negative (positive). When a is very small, there are very few rich individuals. Furthermore, when a is small, the implicit price for the public good for the rich is very large relative to that for the poor. Hence, the amount of public

good provided for the rich is small. This implies that the optimal subsidy from the rich to the poor must be negative in order to prevent the rich from moving to the poor jurisdiction.

As it turns out, if the optimal subsidy chosen by the utilitarian social planner for the federal solution is 0, then it is always better to choose the centralized provision. This means that, in Figure 5, an appropriately extended red curve would never cross the yellow one. We state this formally as follows.

Proposition 3 *Suppose the solution to program (22) is $\hat{s} = 0$. Then the utilitarian social planner prefers the centralized jurisdiction structure to the federal one.*

When the optimal subsidy is zero in the two-jurisdiction structure, there is no redistribution from the rich to the poor. Furthermore, there is duplication in the production of the public good. It is then socially optimal to have only one jurisdiction. The preference for a federal jurisdiction structure over a centralized one can only appear when it is optimal for the central planner in a federal structure to redistribute tax revenues across jurisdictions.

The optimal subsidy can also be examined along the IC_1 curve, where the incentive constraint of the rich is weakly binding. Along this curve, the allocation $sb2$ coincides with the allocation $fb2$. The optimal subsidy is then $\hat{s} = s_1^{fb2}$. As illustrated in Figure 5, there exists a point on the curve above which the two-jurisdiction structure socially dominates the one-jurisdiction structure. This occurs for a value of b that is sufficiently large.

From these two observations, one concludes that the region of social indifference between the federal and the centralized structures is located somewhere between the line for which $\hat{s} = 0$ and that for IC_1 . This suggests that for any given demographic ratio a however small, there exists a high enough interquartile ratio b above which the social planner will always favour a federal structure over a centralized one. Unfortunately, Figure 5 does not provide us with any clue about whether or not this intuition is correct because what happens between the IC_1 curve and the yellow curve does not depend only upon the ratios a and b . It depends upon the four parameters (w_1, w_2, n_1, n_2) . However, as shown in the following lemma, we can prove indirectly that this intuition is correct.

Lemma 2 *For any demographic ratio a such that the constraint IC_1 is binding, there exists a large enough value for the interquartile ratio b at which the utilitarian social planner prefers the two-jurisdiction structure.*

Are the combinations of the demographic and interquartile ratios above which a federal structure is favoured by a utilitarian social planner empirically plausible? Put differently, are the actual (before tax) income distributions observed in - say - developed countries exhibiting such ratios? Table 1 shows the average before tax yearly income (expressed in the currency of the country) in 2010 for each of the 10 deciles of the population. These data are taken from the Luxembourg Income Study data set (see e.g. <http://www.lisdatacenter.org/> for detailed information on this data set).

Table 1: Distribution of before tax income per decile

Decile	Canada (Can \$)	France (€)	Germany (€)	Spain (€)	Denmark (kr)	UK (£)	USA (\$)
1	9730	9410	11040	5086	14574	9083	8026
2	16513	14494	16688	10149	22172	13934	15509
3	21586	17785	21451	13240	26735	17214	20845
4	26612	20766	26090	16137	31861	20497	26372
5	32181	23481	30594	19513	36918	24505	32594
6	37818	26503	35831	23217	41803	28941	39859
7	43650	30065	41486	27283	46972	34471	48497
8	51872	34497	48025	32652	53246	41341	59514
9	63784	41814	58792	40396	62041	51973	76568
10	108164	71578	95160	61618	96228	92659	137153

Applying these data to the (very crude) approximation of the reality that a two-type setting represents requires one to make assumption on the dividing line of the population between "rich" and "poor". It requires also one to treat all incomes on a given side of the dividing line as equivalent (say at their average value). Table 2 below shows the average income among the rich and among poor in the OECD countries of Table 1 for a few "dividing lines" between poor and rich corresponding to as many different assumptions on the demographic ratios. For example, a demographic ratio of 9 corresponds to the situation where the poorest decile of the population is poor and anyone that is not in the poorest decile is rich. At the other extreme a demographic ratio of 1/9 corresponds to the situation where everybody but those who are in the higher decile are poor.

Table 2: Demographic and interquartile ratios

a ratio	b Canada	b France	b Germany	b Spain	b Denmark	b UK	b US
1/9	3.20	2.94	2.95	2.95	2.58	3.45	3.77
1/4	2.87	2.56	2.66	2.77	2.31	3.05	3.40
2/3	2.78	2.37	2.58	2.78	2.23	2.90	3.37
1	2.86	2.38	2.64	2.89	2.27	2.93	3.50
3/2	3.02	2.43	2.74	3.06	2.36	3.01	3.71
4	3.67	2.79	3.22	3.84	2.69	3.38	4.69
9	4.59	3.31	3.77	5.33	3.19	3.98	6.33

As can be seen from this table and Figure 5 above, if one assumes a dividing line rich-poor that generates a demographic ratio in the 3/2-9 range, countries such as Canada, Spain, the UK and the US exhibit a sufficiently large interquartile ratio (above than 3) to justify the choice of a federal structure over a centralized one by a utilitarian social planner. Things are less clear for dividing lines between rich and poor that generate demographic ratios smaller than 1, even though Canada, UK and the US exhibit in this case an interquartile

ratio superior to 3 which could justify the choice of a federal structure. Hence, this table clearly suggests that the ratio combinations for which a federal solution is preferred to the centralized one by a utilitarian social planner using a log-log function are not outlandish.

4 Conclusion

The conclusion of this paper holds in one sentence that can be stated, after Wallace Oates (1972)'s quotation recalled above, as follows. Even when the "cost-savings from the centralized provision of a public good is maximal, the level of (utilitarian) welfare may be at least as high (and typically higher) if Pareto-efficient levels of consumption are provided in each jurisdiction than if any single, uniform level of consumption is maintained." In the model considered in this paper, we have indeed considered the most extreme form of "cost-savings from a centralized provision of a public good" that one can imagine: that of a non-rival public good with no congestion. We have shown in such a setting that if the public authority is imperfectly informed about the willingness to pay of its citizens, it may find optimal to organize the provision of the public good into several distinct jurisdictions rather than in a single one, even at the high cost of provision that results from the replication in several jurisdictions of the cost of providing the very same non-rival public good. The reason for this preference comes from the information that the social planner obtains from having individuals "choosing" their jurisdiction of residence and, therefore, expressing their preferences for their favorite tax and public good packages. In a single "uniform level of consumption" centralized jurisdiction structure, individuals must all pay the same taxes and consume the same amount of public good because under private information, the public authority is incapable of discriminating across types. In a federal structure, the central government may achieve a better targeting of the packages to the tastes of its citizens. Our analysis shows that the benefit of this better targeting may outweigh the cost of unnecessarily replicating the provision of a non-rival public good into several jurisdictions. We have shown more specifically that the superiority of a federal provision of a public good over a centralized one is all the more likely as the heterogeneity in the population is large. Using the somewhat specific case of a two-type population of unequally income individuals, we have shown in particular that federal provision tends to dominate the centralized one when the income differences between the rich and the poor is large, and when the ratio of the rich over the poor is also large. As the fraction of rich people in the population becomes small, the case in favour of a federal solution vanishes even though, for any fraction of rich in the population however how small, it is always possible to find a sufficiently large income discrepancy between the rich and the poor that would make federal provision preferable to the centralized one. We have also suggested that the set of heterogeneity parameters above which the federal solution is preferable to the centralized one is empirically plausible.

While we believe this analysis, and the strong case that it makes in favour of a federal provision of public good, to be of some interest, we are aware of many of its limitations.

For one thing, we have limited our attention to individuals who differ only in contributive capacities (income), and who have the same preferences for the private and the public good. An alternative would have been to consider the case of individuals with the same income, but with different tastes for the public good. We conjecture that one would obtain very similar conclusions to the one obtained here in this case. A more realistic, but analytically much more challenging, situation would have been that where individuals differ both in their income and their preferences.

Another obvious limitation is the restriction of a large part of our analysis to a two-type setting. Yet the difficulty of the problem of the optimal choice of a jurisdiction structure under private information was already significant. Imagine for instance a three-type setting. Then, one would need to consider a large number of possibilities: a centralized solution with all individuals' types pooled into a unique jurisdiction, a "completely decentralized" setting where each type forms a jurisdiction on its own, and two "mixed federal" structures: one where the rich are pooled with the "middle" in one jurisdiction and the poor are left alone and the other where the rich stay alone and the "middle" and the poor form a jurisdiction. Thanks to the single crossing property, these are the only jurisdiction structures that would satisfy the incentive compatibility constraints. But the analysis of all of these cases, with all the varying possibilities for the incentive constraints to bind or not, would have been very tedious.

The last, but not the least, limitation of the analysis is its focus on a specific utilitarian objective for the social planner. Ideally, one would like to generalize the analysis and to study the (second-best) utility possibility sets associated to federalism and centralism so as to identify the set of parameter values for which the frontier of one of the two sets Pareto dominates that of the other. This is a very complicated analysis when incentive constraints bind (see e.g. Maniquet and Sprumont (2010) for the quasi-linear case), and is therefore beyond the scope of this paper.

A Appendix: Proofs

A.1 Proof of Lemma 1

Suppose first that there exists $s_i^* \in [-w_i, +\infty)$ satisfying:

$$z_i^* = z^M(1/n_i, 1, w_i + s_i^*),$$

$$T_i^* = w_i - x^M(1/n_i, 1, w_i + s_i^*)$$

for $i = 1, 2$ such that:

$$n_1 s_1^* + n_2 s_2^* = 0.$$

Yet suppose that, contrary to the statement of the lemma, $(z_1^*, T_1^*, z_2^*, T_2^*)$ is not Pareto-efficient in $B(n_1, n_2, w_1, w_2)$. This means that there exists some allocation $(\hat{z}_1, \hat{T}_1, \hat{z}_2, \hat{T}_2)$ in $B(n_1, n_2, w_1, w_2)$ such that $U(\hat{z}_i, w_i - \hat{T}_i) \geq U(z_i^*, w_i - T_i^*)$ for $i = 1, 2$ with at least one strict inequality. By standard revealed preference arguments, this implies that $\hat{z}_i/n_i + (w_i - \hat{T}_i) \geq z_i^*/n_i + (w_i - T_i^*) = w_i + s_i^*$ for $i = 1, 2$ with at least one strict inequality. Multiplying both sides of each of these inequalities by n_i and summing yields:

$$\begin{aligned} \hat{z}_1 - n_1 \hat{T}_1 + \hat{z}_2 - n_2 \hat{T}_2 &> n_1 s_1^* + n_2 s_2^* \\ &= 0. \end{aligned}$$

But this contradicts the assumption that $(\hat{z}_1, \hat{T}_1, \hat{z}_2, \hat{T}_2)$ is in $B(n_1, n_2, w_1, w_2)$. Conversely, assume that $(z_1^*, T_1^*, z_2^*, T_2^*)$ is Pareto-efficient in the set $B(n_1, n_2, w_1, w_2)$. Define then s_i^* by:

$$s_i^* = \frac{z_i^*}{n_i} - T_i^* \quad (24)$$

for $i = 1, 2$. Since $(z_1^*, T_1^*, z_2^*, T_2^*)$ is Pareto-efficient in $B(n_1, n_2, w_1, w_2)$, it satisfies $n_1 T_1^* + n_2 T_2^* = z_1^* + z_2^*$ so that:

$$n_1 s_1^* + n_2 s_2^* = z_1^* + z_2^* - n_1 T_1^* - n_2 T_2^* = 0$$

as required. Let us now show that $z_i^* = z^M(1/n_i, 1, w_i + s_i^*)$ and $T_i^* = w_i - x^M(1/n_i, 1, w_i + s_i^*)$ for $i = 1, 2$. Thanks to (24), it is clear that, for $i = 1, 2$, $w_i + s_i^* = z_i^*/n_i + w_i - T_i^*$ so that the bundle of public and private good $(z_i^*, w_i - T_i^*)$ satisfies the budget constraint defined by setting the price of public good at $1/n_i$ and the individual income at $w_i + s_i^*$ (using private good as the numéraire). Suppose to the contrary that, for individual 1 (say, but the argument works equally well for individual 2 or for both types of individuals) one has $(z_1^*, w_1 - T_1^*) \neq (z^M(1/n_1, 1, w_1 + s_1^*), x^M(1/n_1, 1, w_1 + s_1^*))$. Since Marshallian demands are functions of the public good prices and income, this means that there exists some bundle (\hat{z}_1, \hat{x}_1) satisfying $w_1 + s_1^* = \hat{z}_1/n_1 + \hat{x}_1$ such that $U(\hat{z}_1, \hat{x}_1) > U(z_1^*, w_1 - T_1^*)$. Defining \hat{T}_1 by $\hat{T}_1 = w_1 - \hat{x}_1$, this means that:

$$n_1 w_1 + n_1 s_1^* + n_2 w_2 + n_2 s_2^* = \hat{z}_1 + n_1(w_1 - \hat{T}_1) + z_2^* + n_2(w_2 - T_2^*).$$

Because $n_1 s_1^* + n_2 s_2^* = 0$, this writes:

$$n_1 \hat{T}_1 + n_2 T_2^* = \hat{z}_1 + z_2^*.$$

But this means that the allocation $(\hat{z}_1, \hat{T}_1, z_2^*, T_2^*) \in B(n_1, n_2, w_1, w_2)$ is Pareto superior to $(z_1^*, T_1^*, z_2^*, T_2^*)$, a contradiction.

A.2 Proof of Proposition 1

Assume that $(z_1^*, T_1^*, z_2^*, T_2^*)$ is efficient in $B(n_1, n_2, w_1, w_2)$ with respect to a scheme of subsidies s_1 and s_2 as per Lemma 1 satisfying $s_1 < 0 < s_2$, and is such that:

$$U(z_2^*, w_2 - T_2^*) \leq U(z_1^*, w_2 - T_1^*). \quad (25)$$

Because the per capita subsidies s_1 and s_2 that support this Pareto-efficient federal provision of public good as per Lemma 1 are progressive, $s_1 = z_1^*/n_1 - T_1^* < 0$ so that $T_1^* > 0$ (since $z_1^* > 0$). Consider then merging jurisdictions 1 and 2 and providing everyone with z_1^* unit of public good while asking a tax payment of T_1^* . If it was feasible, such a centralized provision of the public good would weakly Pareto dominate the federal provision associated with $(z_1^*, T_1^*, z_2^*, T_2^*)$ because type 1 individual would be indifferent and, thanks to inequality (25), type 2 individuals would be weakly better off. Let us show that the centralized provision (z_1^*, T_1^*) is indeed feasible. This amounts to showing that:

$$(n_1 + n_2)T_1^* \geq z_1^*.$$

Since $(z_1^*, T_1^*, z_2^*, T_2^*)$ is feasible and efficient in $B(n_1, n_2, w_1, w_2)$, one has:

$$\begin{aligned} n_1 T_1^* + n_2 T_2^* &= z_1^* + z_2^* \\ &\iff \\ n_1 T_1^* + n_2 T_2^* - z_2^* &= z_1^* \end{aligned}$$

which implies, given that $n_2 T_2^* - z_2^* = -n_2 s_2 < 0$, that $n_1 T_1^* > z_1^*$ and, therefore, that $(n_1 + n_2)T_1^* > z_1^*$ (since $T_1^* > 0$). Because this inequality is strict, it is actually possible to increase public good provision for all without increasing taxes, that is, it is possible to increase the utility level of all.

A.3 Proof of Proposition 2

The proof proceeds as follows. First, we show that:

$$V_R(1/n_1, 1, w_1 - s^*) > \frac{\partial U(\tilde{z}_2, w_2 - \tilde{T}_2)}{\partial x},$$

where $(\tilde{z}_2, \tilde{T}_2)$ denotes the solution of program (11) when the subsidy s is constrained to take the value s^* . We then show that, at the solution of program (11), we must have

$$V_R(1/n_1, 1, w_1 - \hat{s}) < \frac{\partial U(\hat{z}_2, w_2 - \hat{T}_2)}{\partial x}.$$

Finally, we show that, in order to obtain this latter inequality, it must be the case that $\hat{s} \leq s^*$. It is then easy to show that $\hat{z}_2 \leq z_2^M(1/n_2, 1, w_2 + n_1 s^*/n_2)$.

We know that the solution of program (12), s^* , is such that:

$$V_R(1/n_1, 1, w_1 - s^*) = V_R(1/n_2, 1, w_2 + n_1 s^*/n_2),$$

(s^* equalizes the marginal utility of income across individuals). Suppose that we fix s^* and that we solve program (11) by choosing z_2 and T_2 given s^* . The allocation of the rich individual does not

change since the subsidy s^* does not change. The allocation of the poor individual is now implicitly characterized by the two constraints of program (11) that will be binding at the solution of the program (the proof of this is left to the reader):

$$\begin{aligned} n_1 s^* + n_2 \tilde{T}_2 &= \tilde{z}_2 \\ V(1/n_1, 1, w_1 - s^*) &= U(\tilde{z}_2, w_1 - \tilde{T}_2). \end{aligned}$$

In general, there are two allocations that simultaneously satisfy these two equations. However, only one also satisfies the (unwritten) incentive constraint for the poor individual. At this allocation, we have that:

$$\tilde{z}_2 < z_2^M(1/n_2, 1, w_2 + n_1 s^*/n_2)$$

and that:

$$\tilde{T}_2 < T_2^M = \frac{z_2^M(1/n_2, 1, w_2 + n_1 s^*/n_2) - n_1 s^*}{n_2}.$$

Moreover, we know that:

$$V_R(1/n_2, 1, w_2 + n_1 s^*/n_2) = \frac{\partial U(z_2^*, w_2 - T_2^*)}{\partial x}.$$

Let us now show that:

$$\frac{\partial U(z_2^*, w_2 - T_2^*)}{\partial x} > \frac{\partial U(\tilde{z}_2, w_2 - \tilde{T}_2)}{\partial x},$$

that is, let us show that the marginal utility of income of the poor individual decreases when we move from the allocation (z_2^*, T_2^*) to the allocation $(\tilde{z}_2, \tilde{T}_2)$. We first note that we have

$$\begin{aligned} n_1 s^* + n_2 T_2^* &= z_2^* \\ n_1 s^* + n_2 \tilde{T}_2 &= \tilde{z}_2, \end{aligned}$$

that is, both allocations are on the same "budget line" for the poor individual. Let $U_x(n_1 s^* + n_2 T_2, w_2 - T_2)$ be the partial derivative of U with respect to its second argument. Notice that, since $U_{xz} \geq 0$ by super-modularity, one has:

$$\frac{\partial U_x(n_1 s^* + n_2 T_2, w_2 - T_2)}{\partial T_2} = n_2 U_{xz} - U_{xx} > 0. \quad (26)$$

Notice also that, as we move from (z_2^*, T_2^*) to $(\tilde{z}_2, \tilde{T}_2)$ along the budget line, T_2 decreases. Combined with inequality (26), this implies that U_x decreases as well along the budget line. Hence, one has:

$$V_R(1/n_1, 1, w_1 - s^*) > \frac{\partial U(\tilde{z}_2, w_2 - \tilde{T}_2)}{\partial x}.$$

We now show that, at the solution of program (11), we have:

$$V_R(1/n_1, 1, w_1 - \hat{s}) < \frac{\partial U(\hat{z}_2, w_2 - \hat{T}_2)}{\partial x}.$$

We do this by writing down the the first-order conditions of program (11):

$$\begin{aligned} T_2 &: -n_2 U_x(\hat{z}_2, w_2 - \hat{T}_2) + \mu \cdot n_2 + \lambda \cdot U_x(\hat{z}_2, w_1 - \hat{T}_2) = 0 \\ z_2 &: n_2 U_z(\hat{z}_2, w_2 - \hat{T}_2) - \mu - \lambda \cdot U_z(\hat{z}_2, w_1 - \hat{T}_2) = 0 \\ s &: -n_1 V_R(1/n_1, 1, w_1 - \hat{s}) + n_1 \cdot \mu - \lambda \cdot V_R(1/n_1, 1, w_1 - \hat{s}) = 0 \end{aligned}$$

where μ and λ are the multipliers of the two constraints. Substituting for the value of μ drawn from the first condition into the third one yields (after some straightforward manipulations):

$$n_1(U_x(\hat{z}_2, w_2 - \hat{T}_2) - V_R(1/n_1, 1, w_1 - \hat{s})) - \lambda \left(\frac{n_1}{n_2} U_x(\hat{z}_2, w_1 - \hat{T}_2) + V_R(1/n_1, 1, w_1 - \hat{s}) \right) = 0.$$

This can only hold if the first term is positive. Hence,

$$V_R(1/n_1, 1, w_1 - \hat{s}) < \frac{\partial U(\hat{z}_2, w_2 - \hat{T}_2)}{\partial x}.$$

Since the above inequality is reversed at s^* , the final step of the proof requires showing that s must decrease for this inequality to be satisfied. Consider first the effect of s on $V_R(1/n_1, 1, w_1 - s)$. A standard comparative statics exercise shows that, if U is concave, then V is concave in income. This implies that $V_{Rs}(1/n_1, 1, w_1 - s) > 0$. Hence, if s decreases, the marginal utility of income of the rich individual decreases. Consider now the effect of s on $U_x(z_2, w_2 - T_2)$ where (z_2, T_2) are the solution of program (11) for a given s . We know that (z_2, T_2) is at the intersection of the two constraints of program (11). Inserting the budget constraint into the incentive constraint yields

$$V(1/n_1, 1, w_1 - s) - U(n_1 s + n_2 T_2, w_1 - T_2) = 0. \quad (27)$$

This implicitly defines T_2 as a function of s . We want to evaluate the effect of s on $U_x(n_1 s + n_2 T_2, w_2 - T_2)$ when T_2 is characterized by the condition (27). We have:

$$\frac{dU_x(n_1 s + n_2 T_2, w_2 - T_2)}{ds} = U_{xz} \left[n_1 + n_2 \frac{dT_2}{ds} \right] - U_{xx} \frac{dT_2}{ds}, \quad (28)$$

where:

$$\frac{dT_2}{ds} = - \frac{-V_R - n_1 U_z^1}{-n_2 U_z^1 + U_x^1} < 0$$

is calculated along condition (27) and $U^1 = U(n_1 s + n_2 T_2, w_1 - T_2)$ (as opposed to $U = U(n_1 s + n_2 T_2, w_2 - T_2)$). It is easy to show that the denominator of dT_2/ds is negative because of the

single-crossing property and the fact that constraint (27) is binding. If it was positive, then the solution would not be optimal since it would be possible to increase the utility of the poor individual without decreasing that of the rich individual while satisfying all constraints. This implies that $dT_2/ds < 0$.

We now evaluate the sign of the coefficient of U_{xz} in (28).

$$\left[n_1 + n_2 \frac{dT_2}{ds} \right] = \frac{-n_1 n_2 U_z^1 + n_1 U_x^1 + n_2 V_R + n_2 n_1 U_z^1}{-n_2 U_z^1 + U_x^1} = \frac{n_1 U_x^1 + n_2 V_R}{-n_2 U_z^1 + U_x^1} < 0.$$

These sign calculations imply that $dU_x/ds < 0$ in equation (28). Hence, if s decreases, the marginal utility of income of the poor individual increases. We have shown that, at s^* , the marginal utility of the rich individual is higher than that of the poor. We have also shown that it must be lower at the optimal separating allocation. To lower the marginal utility of the rich and increase that of the poor (to reach the optimal separating allocation), we have to decrease s (starting from s^*). Hence, it must be the case that $\hat{s} \leq s^*$. It is easy to show that $\hat{z}_2 \leq z_2^M(1/n_2, 1, w_2 + n_1 s^*/n_2)$. Since the public good is normal, one must have $z_2^M(1/n_2, 1, w_2 + n_1 \hat{s}/n_2) \leq z_2^M(1/n_2, 1, w_2 + n_1 s^*/n_2)$. Hence, handling the incentive constraints of the rich individuals imposes a distortion on the consumption of public good of the poor such that $\hat{z}_2 \leq z_2^M(1/n_2, 1, w_2 + n_1 \hat{s}/n_2)$.

A.4 Proof of Proposition 3

Assume that

$$b = \frac{-\sqrt{1-a} - 2a}{-2a + a\sqrt{1-a}}.$$

Call the solution to program (16) SW^{sb2} .

$$SW^{sb2} = n_1 U(\hat{z}_1, w_1 - \hat{T}_1) + n_2 U(\hat{z}_2, w_2 - \hat{T}_2),$$

where $\hat{z}_1 = n_1 w_1/2$, $\hat{T}_1 = w_1/2$, $\hat{z}_2 = z_2(0)$, $\hat{T}_2 = T_2(0)$. Replacing w_1 by bw_2 and n_1 by an_2 , one can show that:

$$\begin{aligned} \frac{SW^{sb2}}{n_2} = & -a \log[4an_2] + 2a \log \left[-\frac{(2a + \sqrt{1-a})}{\sqrt{1-a} - 2} n_2 w_2 \right] + \log \left[\frac{-1 + \sqrt{1-a} - a}{2a(\sqrt{1-a} - 2)} w_2 \right] + \\ & \log \left[-\frac{(-3+3\sqrt{1-a}+a)(2a+\sqrt{1-a})}{2a(\sqrt{1-a}-2)^2} n_2 w_2 \right] \end{aligned}$$

and that:

$$\frac{SW^{sb1}}{n_2} = a \log \left[\frac{\left(\sqrt{-a^3 - 2(\sqrt{1-a} - 1)a^2 + (2\sqrt{1-a} - 1)a + 1 + a\sqrt{1-a} + a + \sqrt{1-a}} \right) w_2}{2a(2 - \sqrt{1-a})} \right] +$$

$$\log \left[\frac{\left(\sqrt{-a^3 - 2(\sqrt{1-a} - 1)a^2 + (2\sqrt{1-a} - 1)a + 1 - a\sqrt{1-a} + a - \sqrt{1-a}} \right) w_2}{2a(2 - \sqrt{1-a})} \right] +$$

$$(a + 1) \log \left[\frac{-(a+1) \left(\sqrt{-a^3 - 2(\sqrt{1-a} - 1)a^2 + (2\sqrt{1-a} - 1)a + 1 + a\sqrt{1-a} - 3a - \sqrt{1-a}} \right) n_2 w_2}{2a(2 - \sqrt{1-a})} \right].$$

Thanks to *Mathematica*, we compute the difference $SW^{sb2}/n_2 - SW^{sb1}/n_2$. We can show that the resulting expression is independent of w_2 and n_2 . We then set without loss of generality $w_2 = n_2 = 1$. This implies that the difference in welfare only depends on the ratio a . We can plot it over the interval $a \in]0, 1]$. The graph demonstrates that the expression is negative over this domain of a . Consequently, the one-jurisdiction structure yields higher social welfare than the two-jurisdiction structure does.

A.5 Proof of Lemma 2

One needs to compute the difference $SW^{sb2} - SW^{sb1}$. The difficulty lies in characterizing the optimal level of subsidy \hat{s} in the allocation $sb2$ where the IC1 constraint binds. Instead of doing so, we compute this difference using the subsidy s_i^{fb2} . Since this subsidy is suboptimal when IC_1 binds, this computation underestimates the difference $SW^{sb2} - SW^{sb1}$. Using *Mathematica*, we get:

$$SW_f^{sb2}/n_2 - SW^{sb1}/n_2 = 2a \log \left[\frac{an_2 w_2 (ab+1)}{a+1} \right] - a \log[4an_2] +$$

$$\log \left[\frac{n_2 w_2 \left(\sqrt{a^3(-b^2) + a^2(4b^2 - 6b + 1) + a(4b^2 - 2b - 1) + b^2 + a - b + 2} \right) \left(\sqrt{a^3(-b^2) + a^2(4b^2 - 6b + 1) + a(4b^2 - 2b - 1) + b^2 - 2ab + a - b} \right)}{(a+1) \left(-\sqrt{(a(b+2) + 2b + 1)^2 - 8b(a+1)^2 + ab - 2a + 2b - 3} \right)} \right] -$$

$$(a + 1) \log \left[\frac{1}{4} n_2 w_2 \left(-\sqrt{(a(b+2) + 2b + 1)^2 - 8b(a+1)^2 + ab + 2a + 2b + 1} \right) \right] -$$

$$a \log \left[\frac{w_2 \left(\sqrt{(a(b+2) + 2b + 1)^2 - 8b(a+1)^2 + 3ab - 2a + 2b - 1} \right)}{4(a+1)} \right]$$

where SW_f^{sb2} refers to the second-best social welfare with the suboptimal subsidy s_i^{fb2} . We can show that the sign of this expression only depends on the ratios a and b , that is, it is independent of n_2 and w_2 . We then fix the ratio a and take the limit of $SW_f^{sb2} - SW^{sb1}$ when b goes to infinity. We can show that this difference converges to infinity. This implies that, for any demographic ratio a , there exists a high enough interquartile ratio b such that the two-jurisdiction structure is socially preferable

References

- ARONSSON, T., AND S. BLOMQUIST (2008): “Redistribution and Provision of Public Goods in an Economic Federation,” *Journal of Public Economic Theory*, 10, 125–143.
- BOADWAY, R., AND F. FLATTERS (1982): “Efficiency and equalization payments in a federal system of government: A synthesis and extension of recent results,” *Canadian Journal of Economics*, 15, 613–633.

- BORDIGNON, M., P. MANASSE, AND G. TABELLINI (2001): "Optimal Redistribution under Asymmetric Information," *American Economic Review*, 91, 709–723.
- BREUILLÉ, M.-L., AND R. J. GARY-BOBO (2007): "Sharing Budgetary Austerity under Free Mobility and Asymmetric Information : an Optimal Regulation Approach to Fiscal Federalism," *Journal of Public Economics*, 91, 1177–1196.
- BUCHANAN, J. M. (1950): "Federalism and Fiscal Equity," *American Economic Review*, 40, 583–599.
- CORNES, R. C., AND E. C. SILVA (2000): "Local Public Goods, Risk Sharing and Private Information in Federal Systems," *Journal of Urban Economics*, 47, 39–60.
- (2002): "Local Public Goods, Interegional Transfers and Private Information," *European Economic Review*, 46, 329–356.
- FLATTERS, F., V. HENDERSON, AND P. MIESZCOWSKI (1974): "Public goods, efficiency, and regional fiscal equalization," *Journal of Public Economics*, 3, 99–112.
- GRAVEL, N., AND M. POITEVIN (2006): "The Progressivity of Equalization Payments in Federations," *Journal of Public Economics*, 90, 1725–1743.
- GREENBERG, J. (1983): "Local Public Goods with Mobility: Existence and Optimality of a General Equilibrium," *Journal of Economic Theory*, 16, 137–150.
- MANIQUET, F., AND Y. SPRUMONT (2010): "Sharing the cost of a public good: an incentive-constrained axiomatic approach," *Games and Economic Behavior*, 68, 275–302.
- MIRRLEES, J. A. (1971): "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175–208.
- (1976): "Optimal Tax Theory: A Synthesis," *Journal of Public Economics*, 6, 327–358.
- (1986): "The Theory of Optimal Income Taxation," in *Handbook of Mathematical Economics*, ed. by K. J. Arrow, and M. D. Intriligator. North Holland, Amsterdam.
- OATES, W. E. (1972): *Fiscal Federalism*. Harcourt Brace Jovanovich, New York.
- (1999): "An Essay on Fiscal Federalism," *Journal of the Economic Literature*, 37, 1120–1149.
- TIEBOUT, C. M. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64, 416–424.
- WILDASIN, A. (1987): "Theoretical Analysis of Local Public Economics," in *Handbook of Regional and Urban Economics*, ed. by E. Mills. North Holland.